HUMAN ERROR IN MARITIME OPERATIONS: ANALYSES OF ACCIDENT REPORTS USING THE LEXIMANCER TOOL

Michelle R. Grech, Tim Horberry and Andrew Smith
Key Centre of Human Factors and Applied Cognitive Psychology,
University of Queensland,
Brisbane, Qld

This paper focuses on the problem of lack of Situation Awareness (SA) by mariners. An analysis of a large number of accident reports was conducted in order to determine the extent to which SA is a relevant issue in merchant shipping operations. For the first time use was made of the *Leximancer* tool due to its ability to rapidly analyse large amounts of textual information. One major function of this research was to examine the accuracy and usefulness of such a data analysis tool by comparing the results of this computer analysis with that of a 'manual' analysis (performed by two raters).

Our results underline the importance of SA in decision-making processes in the maritime domain: a large number of investigated maritime accidents were partly due to loss of SA. Also, the results of the *Leximancer* tool were found to be comparable to the manual analysis, thus suggesting further use of such a system for accident report analysis in other transportation domains.

INTRODUCTION

A number of studies conducted by various maritime organizations reported more than 75 % of accidents of ships worldwide are due to human and organisational errors (International Maritime Organization, 1994). Hence, any attempts to reduce accidents at sea should concentrate on eliminating errors on board ships, since this is where the problem is greatest and where the biggest improvements should be made.

One way to do this is the analysis of accident report forms, to attempt to uncover causal factors. However, this approach is problematic - in the maritime domain analysis of accident reports are sometimes subject to intense analyses depending on the nature and severity of the accident. This can be a very time consuming task (Caridis, 1999). Additionally, there is no standardisation in maritime accident reporting forms across the world (European Commission, 2001 p. 19).

In this paper we focus particularly on one aspect of human error – lack of Situation Awareness (SA) - and discuss the way in which the concept can be applied to the maritime domain. To the present day research on SA has primarily been restricted to the aviation and recently to a certain extent to the medical domain (Endsley, 1996). However, Endsley (1996) expresses the view that SA is equally important in other complex and dynamic environments, such as the maritime domain. A review of the literature clearly indicates that previously very little work has been carried out in the maritime domain on issues related to SA (Grech & Horberry, 2002).

In this paper we will outline the importance of SA in operator decision-making processes within the maritime domain. To explore the problem we utilised a modified version of a multilevel SA taxonomy developed by Endsley (1995), shown in Table 1. Since it is difficult to reliably discriminate between the sub-levels, we decided to combine them, and collapse the taxonomy into 3 levels as shown in Table 1. Simplifying Endsley's taxonomy in this way made it more fitting for our purposes. Hence the sub level 'memory loss' which was initially under SA level 1 was transferred to SA level 2. Another reason for modifying this taxonomy was the ease of adaptability for use by the Leximancer tool. Furthermore it was more suitable for statistical analyses.

In essence, this model divides SA occurrences into shortcomings of the cognitive psychology paradigm of perception, cognition, and projection of future events. Such a type of model has previously been used to study SA related problems in the aviation domain (Jones & Endsley, 1996).

Table 1 Endlsey's modified SA error taxonomy

Level 1: Failure to correctly perceive information such as;	
o	Data not available
o	Data hard to discriminate or detect
o	Failure to monitor or observe data
o	Misperception of data
Level 2: Failure to correctly integrate or comprehend information such as;	
o	Lack of poor mental model
o	Use of incorrect model
o	Over reliance on default values
o	Memory loss
Level 3: Failure to project future actions or state of the system such as;	
o	Overprojection of current trends

The purpose of this research was twofold: first to analyse and quantify the problem of lack of SA in maritime accidents, second, to measure the efficacy and accuracy of the Leximancer tool for accident report analysis, with a view to using such tools for analysis of a wide range of accident and incident reports.

LEXIMANCER

In this study we made use of the Leximancer tool due to its ability for rapidly mapping text documents using a set of conceptual dimensions. Leximancer is a new software system for performing conceptual analysis of text data. It provides both automatic analyses (through machine learning) and also customised content analyses using defined concept classifiers. This offers an efficient way of quantifying and exploring large text documents using a classification system of learned lexical concepts, rather than just keywords. The operational method employed by the Leximancer involves the following steps:

1. Text preparation: Standard techniques are employed, including name and term preservation;
2. Creation of concept classifiers: These can be devised in collaboration with a domain expert. Otherwise concepts can be selected automatically using a novel algorithm for finding significant seed words to reflect the themes present in the data;
3. Learning the concept classifiers: A machine learning algorithm is used to find the optimal thesaurus words from the text data;
4. Classification: Text is classified using these concepts, to a defined sentence resolution;
5. Indexing: The resulting tagged text is indexed to the required resolution using the entities and properties;
6. Mapping: Entity concepts are clustered according to weight and relationship, to create a concept cluster map;
7. User interface: A simple hypertext browser is used for exploring the classification system in depth.

METHOD

Maritime Accident Reports

A total of 177 maritime accident reports (public domain) originating from eight different countries were analysed. These were chosen to include a wide selection of different types of accidents. Maritime accidents occurring between the years 1987 and 2001 were analysed. Vessels' year of built ranged between 1952 and 2000.

Procedure

In this research, the procedure was undertaken in three separate stages.

Stage 1: Data Exploration: Initial Analyses with Leximancer.

To gain a rapid insight into the content of the data set, Leximancer was configured in a way to map the document set in a fully automatic mode. This means that the Leximancer was allowed to analyse the data freely without any human intervention. An automatic unsupervised map was thus produced. Although we were satisfied with the successful outcome of this first stage analyses technique, most of the links obtained between the maritime concepts extracted from this initial study were however quite obvious for domain experts.

The second step of the analyses involved the construction of customised concept keywords to project the text data onto a set of measurable dimensions. Hence, thesaurus-like concepts of interest were manually specified. Due to a limited mention of human factors in maritime safety reports and the importance of humans in being the cause of most accidents, the concept dimensions were designed to capture various accident events and human factor causes, wherever possible. Design of these keywords for more abstract or complex concepts required some care and progressive refinement, but the use of machine learning facility required that only a few of the highly relevant words required manual selection. Furthermore, the keywords only need to be selected once, since machine learning adopts the vocabulary automatically to new document sets. The aim of the second stage of this analyses were to;

1. try to observe whether the Leximancer tool was able to pick up concepts which are not frequently used in these reports;
2. analyse whether it was capable of tracing their associated links; and
3. analyse its usability as a data exploration tool.

A heuristic analysis of this by the project team confirmed that the Leximancer can be utilised to analyse defined concepts in maritime casualty reports. Therefore, it was decided to go one stage further and focus on SA problems in the maritime domain.

Stage 2: Manual Coding.

Using the three-level SA error taxonomy (table 1), a pilot study was conducted on a sub section of 26 of the 177 maritime reports. In order to reduce subjective bias as much as possible two persons were used during the manual coding process to increase reliability of the study. Ratings were conducted independently. One was a domain expert, the other a human factors specialist. During this pilot study hand coding of the reports was conducted in the 3 SA levels. This coding was used as a learning tool for the Leximancer and also to compare data to Leximancer results.

An example of a manually coded segment from one of the reports follows:

> *'From about 0645 to 0715 the mate had become preoccupied with arranging and making private telephone calls while the ship was in cellular phone range of the coast, rather than monitoring the ship's course, speed, position and his other watchkeeping duties* ***(SA1)****'*
> (Australian Transport Safety Bureau, 2001).

Stage 3: Leximancer Coding.

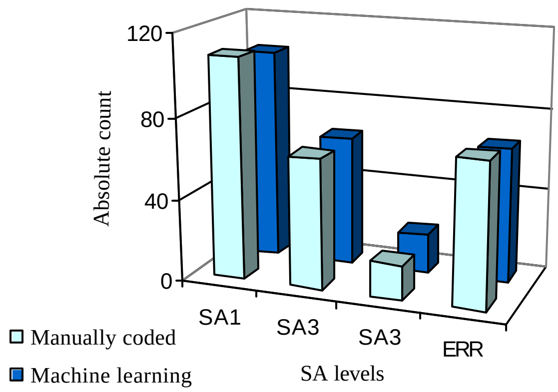
Leximancer analyses on selected reports using machine learning. The Leximancer was allowed to conduct supervised training and classification on our sub-set of 26 reports, which were initially manually coded. Through this we sought to verify the accuracy of the Leximancer analyses as compared to the manually coded results. Prior to the Leximancer analyses all manually coded events were removed from the reports in order to allow the Leximancer to analyse the data without 'prevarication'. Furthermore, in order to eliminate as much as possible 'noise' and 'distraction' from the reports, some text that we considered 'irrelevant', such as replicated and unrelated text that did not refer to the actual cause of the accident was removed. This decreased the length of the reports and contributed to far more accurate results. A three sentences training and classification block was used for machine learning.

Leximancer analyses on whole data set using machine learning. Following this analysis, the Leximancer was than allowed to analyse the rest of the 151 maritime reports by means of the machine learned classifier it acquired from the manually coded reports. Although machine learning on the manually coded reports was conducted on three sentence blocks, analyses of the rest of the data set was based on one sentence blocks.

RESULTS

The results from the manual coding stage revealed that 71% of human errors were SA related problems. Of the SA errors identified during the manual coding process, 58.5% were level 1 SA errors, 32.7% were level 2, and 8.8% were level 3 (Grech & Horberry, 2002). A comparison between the results of manual coding and the Leximancer on these 26 reports revealed a very close proximity between hit rates as shown in figure 1.

Figure 1 Comparison between SA absolute counts using manual coding and Leximancer, on abridged selected reports.



The result ERR in figure 1 is an indication of other human errors encountered which are not SA related problems.

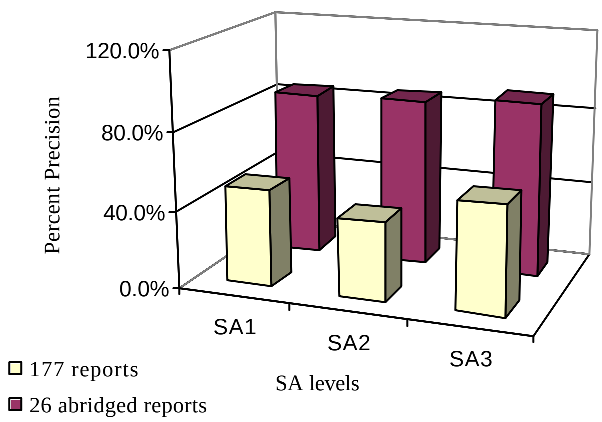
In order to ensure that the comparison of data provides statistically significant results a paired sample t-test was conducted to test for differences between the manual coding and Leximancer absolute counts – no significant difference was found ($p > 0.05$).

In all the 3 SA levels, the percentage precision was more than 84%. Precision is a measure of the percentage of tagged text retrieved by Leximancer which are relevant.
Another measure that was taken into account is recall, which is a percentage measure of those relevant, which are retrieved. Recall provides an indication of the percentage of events (tagged text) selected during manual coding, which were also singled out by the Leximancer. The total percentage recall for all the three SA levels was 89%. Both precision and recall provide an indication of the accuracy of the Leximancer in analysing these reports using the manually coded reports as benchmark. In the field of automatic text classification, the precision and recall obtained in this study are good benchmark results considering that the unit classification is a three-sentence segment. Nevertheless, it should be noted that the fact that we used a relatively small sample size, the selected reports where abridged, and we classified on the text we trained upon in this analyses, did contribute to some extent to these positive results.

As expected precision degraded substantially when the larger data set was analysed. One main reason for this is that the larger data set was not abridged. Analyses of the entire data set (177 maritime accident reports) using machine learning based on one sentence block classification produced relatively good precision (48%). Calculation of this was based on a batch sample of 60 location entities taken for all three SA levels and calculating the percentage precision. Recall could not be measured in this case, as it is very hard to calculate for such a large data set.

A comparison of percentage precision between the Leximancer analyses using machine learning on the sub-set of the reports and the whole data set is shown in figure 2.

Figure 2 Comparison between percent precision using machine learning on abridged, and whole data set.



DISCUSSION

The initial analysis undertaken in this study has shown that a large proportion of human error problems in the maritime domain can be grouped into a single category that can be labeled "loss of SA". The distribution of SA- errors obtained in this study was also comparable to previous studies conducted in the aviation domain. Almost identical percentages were obtained between the manual coding and with the Leximancer analyses. This provides good evidence that problems with situation awareness are indeed a primary factor underlying maritime casualties. Additionally, analysis of accident report forms can be successfully undertaken by using a tool such as the Leximancer.

It should be noted however that a number of problems related to the casualty reports used could cause some bias in our research. A number of factors concerned with the maritime accident reports utilised during this study possibly contributed to variations in the results obtained.

- o Incident reports, independent from accident reports, would have been very useful in this study. It has been indicated that there are generally about 100 plus incidents and 10 to 100 near misses to every accident. (Bea, 1999). Van der Schaaf (1991) suggests that the analyses of near misses provides the link between highly visible and detectable (but rare) accidents and very frequent, but almost invisible, potentially dangerous behavioural acts. Hence, confining our analyses to accident reports only and not to incident and near miss reports or even to normal operations might have resulted in other losses of SA not being captured.
- o Accident investigation procedures across countries differ widely due to the various methods and procedures utilised, even though they do have certain common features.
- o Apart from the modified sub-set most of the remaining reports contained repetitive text and other text not related to the actual accident reports which might have distorted the Leximancer results, hence explaining the substantial degradation in precision when the whole subset was analysed.
- o A major 'technical' problem creating bias in the results is that the information contained in maritime accident reports is not consistent. Detailed accident investigations have often led to efforts at preventing the repetition of accidents that have already occurred through the identification of immediate causes at the expense of the underlying causes. This has happened mainly due to technical investigators being experts in the area of proximate cause, but having minimal or no knowledge of human factors in accident causation and safety management. Work conducted by the European Commission on this aspect indicates that this issue makes it almost impossible to establish relationships between rare casualties and more frequent incidents, and normal operations (European Commission, 2001).

The following issues are considered to merit further attention in future considerations;

1. Caridis (1999) states that by introducing a common approach to marine casualty investigations and the reporting of such events, the international maritime community may become better informed on factors that cause, or contribute to, accidents at sea. One of the most critical factors in achieving this goal is to establish a common format for reports to facilitate the publication and sharing of lessons learnt. In actual fact work has already started to implement such a procedure within the EU (European Commission, 2001).
2. The role of advanced technology systems aimed at reducing the risk of accidents should be further investigated. Advanced technology can be viewed as beneficial for operators in terms of increased information processing power However, a study conducted by Grech & Horberry (2002) revealed that one of the consequences of increasing technology levels is a loss of situation awareness, which can significantly affect performance in abnormal, time-critical circumstances. The decision to rely, or not to rely on automation can be one of the most important decisions each member of the crew has to make.
   In some cases, crew members may rely too much on the automation and fail to check and monitor its performance. On the other hand, they may have lost confidence in their own activities to perform the tasks manually, which obviously defeats the purpose of automation in itself. Following these results there should be growing concern within the maritime domain regarding potential negative effects of automation on performance of shipboard tasks. This is the direction for future research at the University of Queensland, Australia.

Despite the above mentioned qualifications, this research has demonstrated that lack of SA is a serious problem in the maritime domain. The suggested countermeasures could help tackle this problem. Furthermore, initial results have demonstrated that the Leximancer is a valuable tool (in terms of speed, accuracy and revealing unexpected trends) to help analyse such large data sets.

REFERENCES

Australian Transport Safety Bureau (2001). Navigation (Marine Casualty) Regulations- Investigation into the grounding of the Malaysian flag container ship Bunga Teratai Satu on Sudbury Reef, Great Barrier Reef on 2 November 2000 (162). Canberra, Marine Investigation Unit (MIU).

Bea, R. (1999). Risk Assessment & Management (RAM) of Marine Systems: Past, Present, Future. Paper presented at the Proceedings of the State of the Art Pipeline Risk Management Conference, Perth, Western Australia.

Caridis, P. (1999). Casualty Analysis Methodology for Maritime Operations (Final Report for Publication C01.FR.003). Athens.

Endsley, M. R. (1995). Measurement of situation awareness in dynamic systems. Human Factors, 37(1), 65-84.

Endsley, M. R. (1996). Situation awareness measurement in test and evaluation. In T. G. C. O'Brien, Samuel G. (Ed.), Handbook of human factors testing and evaluation. (pp. 159-180). Hillsdale, NJ, US: Lawrence Erlbaum Associates.

European Commission. (2001). Reports selected methods suitable to cover end-user needs -Suggested form for human related data collection during accident investigation (WP 4, Deliverable D4.2, Restricted). Brussels.

Grech, M., & Horberry, T. (2002). Human Error in Maritime Operations: Situation Awareness and Accident Reports. Paper to be presented at the 5th International Workshop on Human Error, Safety and Systems Development, Newcastle, Australia.

International Maritime Organization. (1994). Better standards, training and certification: IMO's response to human error. IMO news.

Jones, D. G., & Endsley, M. R. (1996). Sources of situation awareness errors in aviation. Aviation, Space, and Environmental Medicine, 67(6), 507-512.

Van der Schaaf, T. W. (1991). A Framework for Designing Near Miss Management Systems. In T. W. Van der Schaaf & D. A. Lucas & A. R. Hale (Eds.), Near Miss Reporting as a Safety Tool. Oxford.